

# CARMA Processes driven by Non-Gaussian Noise

Robert Stelzer<sup>\*</sup>

We present an outline of the theory of certain Lévy-driven, multivariate stochastic processes, where the processes are represented by rational transfer functions (Continuous-time AutoRegressive Moving Average or CARMA models) and their applications in non-Gaussian time series modelling. We discuss in detail their definition, their spectral representation, the equivalence to linear state space models and further properties like the second order structure and the tail behaviour under a heavy-tailed input. Furthermore, we study the estimation of the parameters using quasi-maximum likelihood estimates for the auto-regressive and moving average parameters, as well as how to estimate the driving Lévy process.

## 1 Introduction

In many applications an observer (scientist, engineer, analyst) is confronted with series of data originating from one or more physical variables of interest over time. Thus, he has an observed (multivariate) time series and will often either be interested in removing (measurement) noise to extract the signal more clearly or in modelling the observed process, including its random components.

In both situations stochastic models may very well be appropriate. This is clear when one is mainly interested in removing noise, but when intending to model the observed value it is also very often appropriate to enrich a physical model by a random component to capture fluctuations and shortcomings of the physical model. The driving stochastic process (the “noise”) may have interest on its own (as is the case with economic models), or it has to be modelled well to extract the interesting information as well as possible (e.g., as is common practice in telecommunication links)

---

<sup>\*</sup>Institute of Mathematical Finance, Ulm University, Helmholtzstraße 18, D-89069 Ulm, Germany. *Email:* robert.stelzer@uni-ulm.de, <http://www.uni-ulm.de/mawi/finmath>

The easiest way to obtain a model with randomness for the variables of interest would be to assume that all observed values are independent and identically distributed (iid) random variables or that they follow a physical model plus iid noise. However, in most series observed consecutive values are heavily dependent and thus more sophisticated models are needed. A flexible but at the same time very tractable class of models is given by linear random processes. In the discrete time setting these models are well-known as autoregressive moving average (ARMA) processes and they are given in terms of a general order linear difference equation where an iid noisy input sequence introduces all randomness. The latter is also referred to as linear filtering of a white noise.

In many situations it is more appropriate to specify a model in continuous time rather than in discrete time. These include high-frequency data, irregularly spaced data, missing observations or situations when estimation and inference at various frequencies is to be carried out. Moreover, many physical models are formulated in continuous time and, hence, such an approach is often more natural.

In the following we consider linear random processes in continuous time, referred to as continuous time autoregressive moving average (CARMA) processes. Intuitively, they are given as the solution to a higher order system of linear differential equations with a stochastic process as the input, which can be seen as linearly filtering the random input.

One important question is which random input to take in the continuous time set-up. Clearly, the random process should correspond in some sense to the idea of white noise. Understanding the latter in the strict sense means using independent increments, in the weak sense it means uncorrelated increments and so the variance has to be finite. Recall that for random variables uncorrelatedness is equivalent to independence only if the random variables are Gaussian, i.e. they have a normal distribution. A linear random process driven by Gaussian white noise has again Gaussian distributions. However, in many situations it is not appropriate to assume Gaussianity of the variables of interest, since the observed time series often exhibit features like skewness or heavy-tails (i.e. very high or low values are far more likely to occur than in the Gaussian setting), which contradict the Gaussian assumption. Demanding uncorrelated but not necessarily independent increments does not lead to a nice class of processes nor to nice theoretical results.

Hence, a good modelling strategy where the resulting process is reasonably tractable and the driving process' probability distribution is allowed to have "fat tails" is to demand that the random input shall have independent as well as stationary increments, i.e. increments over time intervals of the same length have the same distribution. They then have a time homogeneity feature and resemble the iid noise of the discrete time set-up. The resulting class of possible driving processes are the so-called Lévy processes, which have been studied in detail and form a both highly versatile and highly tractable family. An interesting feature is that linear processes driven by general Lévy processes may exhibit jumps and thus allow the modelling of abrupt changes, whereas Gaussian linear processes have continuous sample paths.

In the remainder of this paper we proceed as follows. First, we introduce Lévy processes in detail. Thereafter, we give a proper definition of CARMA processes, discuss their relation to linear filtering via a stochastic Fourier (spectral) representation and summarize central properties of CARMA processes. Next, we briefly explain the equivalence to linear state space models and the relation to stochastic control and signal processing. Finally, we discuss the statistical

estimation of the parameters and the underlying Lévy process and conclude with some additional remarks.

Throughout we will focus on developing the main ideas for CARMA processes. For more mathematical details as well as comprehensive references we refer the interested reader to the original literature especially the works [2], [8–10], [15–17], [12], [13], [42] and [55]. For a historic perspective the monograph [47] may be interesting as well as [19] which is the first paper where Gaussian CARMA processes appeared under the name of Gaussian processes with rational spectral density.

## 2 Lévy processes

A Lévy process  $L = (L_t)_{t \in \mathbb{R}^+}$  is a stochastic process with independent and stationary increments. In the following we consider only Lévy processes taking values in the  $m$ -dimensional vector space  $\mathbb{R}^m$  (with  $\mathbb{R}$  the real numbers and  $m$  some positive integer). Note that a stochastic process  $(X_t)_{t \in \mathbb{R}^+}$  can be either seen as a family of random variables indexed by the positive real numbers  $\mathbb{R}^+$  or as a random function mapping the positive real numbers to  $\mathbb{R}^m$ . More precisely we have the following definition:

**Definition 2.1.** An  $\mathbb{R}^m$ -valued stochastic process  $L = (L_t)_{t \in \mathbb{R}^+}$  is called Lévy process if

- $L_0 = 0$  a.s.,
- $L_{t_2} - L_{t_1}, L_{t_3} - L_{t_2}, \dots, L_{t_n} - L_{t_{n-1}}$  are independent for all  $n \in \mathbb{N}$  and  $t_1, t_2, \dots, t_n \in \mathbb{R}^+$  with  $0 \leq t_1 < t_2 < \dots < t_n$ ,
- $L_{t+h} - L_t \stackrel{\mathcal{D}}{=} L_{s+h} - L_s$  for all  $s, t, h \in \mathbb{R}^+$  (“ $\stackrel{\mathcal{D}}{=}$ ” denoting equality in distribution),
- $L$  is continuous in probability, i.e. for all  $s \in \mathbb{R}^+$  we have  $L_t - L_s \xrightarrow{P} 0$  as  $t \rightarrow s$ .

It can be shown (cf. [53] for a detailed proof) that the class of Lévy processes can be characterized fully at the level of “characteristic functions”, which we now introduce. Let  $\langle \cdot, \cdot \rangle$  indicate the natural inner product in  $\mathbb{R}^m$  and  $X$  is an  $\mathbb{R}^m$ -valued random variable, then its *characteristic function* is defined as  $\psi_X(u) = E(e^{i\langle u, X \rangle})$ . The characteristic function of a Lévy process can always be represented in the *Lévy-Khintchine form*

$$E\left(e^{i\langle u, L_t \rangle}\right) = \exp\{t\psi_L(u)\}, \quad \forall t \geq 0, u \in \mathbb{R}^m, \quad (2.1)$$

with

$$\psi_L(u) = i\langle \gamma, u \rangle - \frac{1}{2}\langle u, \Sigma_G u \rangle + \int_{\mathbb{R}^m} (e^{i\langle u, x \rangle} - 1 - i\langle u, x \rangle 1_{[0,1]}(\|x\|)) v(dx), \quad (2.2)$$

where  $\gamma \in \mathbb{R}^m$ ,  $\Sigma_G$  is a  $m \times m$  positive semi-definite matrix and  $v$  is a measure on  $\mathbb{R}^m$  that satisfies  $v(\{0\}) = 0$  and  $\int_{\mathbb{R}^m} (\|x\|^2 \wedge 1) v(dx) < \infty$ . The measure  $v$  is referred to as the Lévy measure of  $L$ .

and  $\|x\|^2 \wedge 1$  is short for  $\min\{\|x\|^2, 1\}$ . Finally,  $1_A(x)$  generically denotes the indicator function of a set  $A$ , i.e. the function which is one if  $x$  is an element of  $A$  and zero otherwise. Together  $(\gamma, \Sigma_G, v)$  are referred to as the characteristic triplet of  $L$ .

Regarding the paths of a Lévy process, i.e. the “curve of  $L$  as a function of time  $t$ , it can be shown that without loss of generality, a Lévy process may be assumed to be right continuous and have left limits.

It should be noted that many well-known stochastic processes are Lévy processes. Examples are Brownian motion, also referred to as the Wiener process or “Gaussian white noise”, the Poisson process, which has jumps of size one and remains constant in between the jumps, which occur after iid exponentially distributed waiting times, and  $\alpha$ -stable Lévy motions, sometimes called Lévy flights. Compound Poisson processes are Poisson processes where the fixed jump size one is replaced by random iid jump sizes independent of the interarrival times of the jumps. It can be shown that all Lévy processes arise as limits of such compound Poisson processes.

A better understanding of what Lévy processes really are is provided by the Lévy-Itô decomposition of their paths. It states that a Lévy process is the sum of the deterministic linear function  $\gamma t$ , a Brownian motion with covariance matrix  $\Sigma_G$ , the sum of the big jumps which form a compound Poisson process and the compensated sum of the small jumps (i.e. the sum of the small jumps minus their expected value). The quantity  $v(A)$  gives for any measurable set  $A \subset \mathbb{R}^m$  the expected number of jumps with size in  $A$  occurring in a time interval of length one. In Figure 1 a univariate Lévy process which is the sum of the linear function  $t$ , in this case with  $\gamma = 2$ , a standard Brownian motion, with  $\Sigma_G = 1$ , and a Poisson process, with  $v(\{1\}) = 1$ ,  $v(\mathbb{R} \setminus \{1\}) = 0$  is depicted together with its individual components.

Whenever  $\int_{\mathbb{R}^m} (\|x\| \wedge 1) v(dx) < \infty$ , we can replace the compensated sum of small jumps simply by the sum of the small jumps adjusting also the slope of the deterministic component. We have actually already done this in Figure 1 where the resulting slope of the deterministic function is  $\gamma - \int_{\mathbb{R}} x v(dx) = 1$ . If  $v(\mathbb{R}) < \infty$ , we have finitely many jumps in any bounded time interval and the jumps form actually a compound Poisson process. Otherwise, we have infinitely, but countably many jumps in any bounded time interval. The reason why we have in general a component referred to as “the compensated sum of the jumps” (i.e., it results from a certain limiting procedure see e.g. [53]) is that in general the jumps are not summable. This is equivalent to the fact that the paths have infinite variation, like Brownian motion. Infinite variation intuitively means that the curve described by the stochastic process over finite time intervals has an infinite length. Clearly, this means that the fluctuations of the process over small time intervals are rather vivid.

In Figures 2 and 3 you can see simulations of different pure jump Lévy processes, i.e. in these cases  $\gamma = 0$  and  $\Sigma = 0$ . So there is neither a deterministic drift nor a Brownian motion present. All these processes have infinite activity, i.e. infinitely many jumps in any time interval. Figure 2 depicts a so-called normal inverse Gaussian Lévy process which has heavier tails than a Brownian motion, but still is rather tame, because it has finite moments of all orders, i.e.  $E(|L_t|^r) < \infty$  for all  $t, r \in \mathbb{R}^+$  and also some exponential moments. In contrast to this the stable processes of Figure 3 have very heavy tails, because they do not have a finite variance and the 0.5-stable processes does not even have a finite mean. Whereas the NIG and 1.5-stable processes have infinite variation, the small jumps of the 0.5-stable Lévy process are summable.

Most of the time we will work with Lévy processes defined on the whole real line, i.e. in-

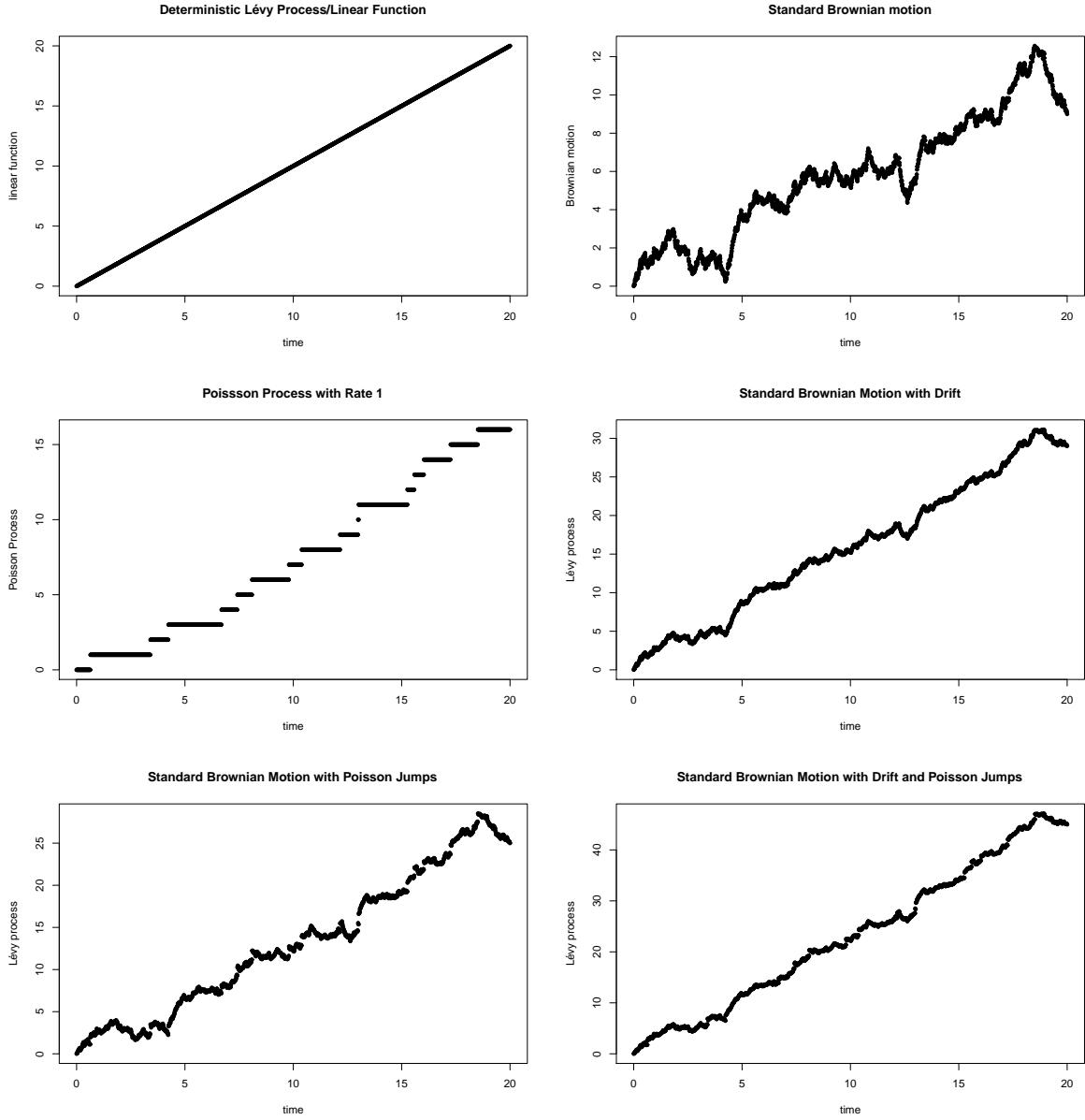


Figure 1: A Lévy process and its components: The complete Lévy process is depicted in the lower right display. In the upper row the deterministic drift component is depicted on the left and the standard Brownian motion component on the right. The left display in the middle row shows the standard (rate one) Poisson component and the right one the Brownian motion and the deterministic component added together. In the last row on the left the Brownian component plus the Poisson jumps are depicted. Note that the scaling of the y-axis is different in the individual plots.

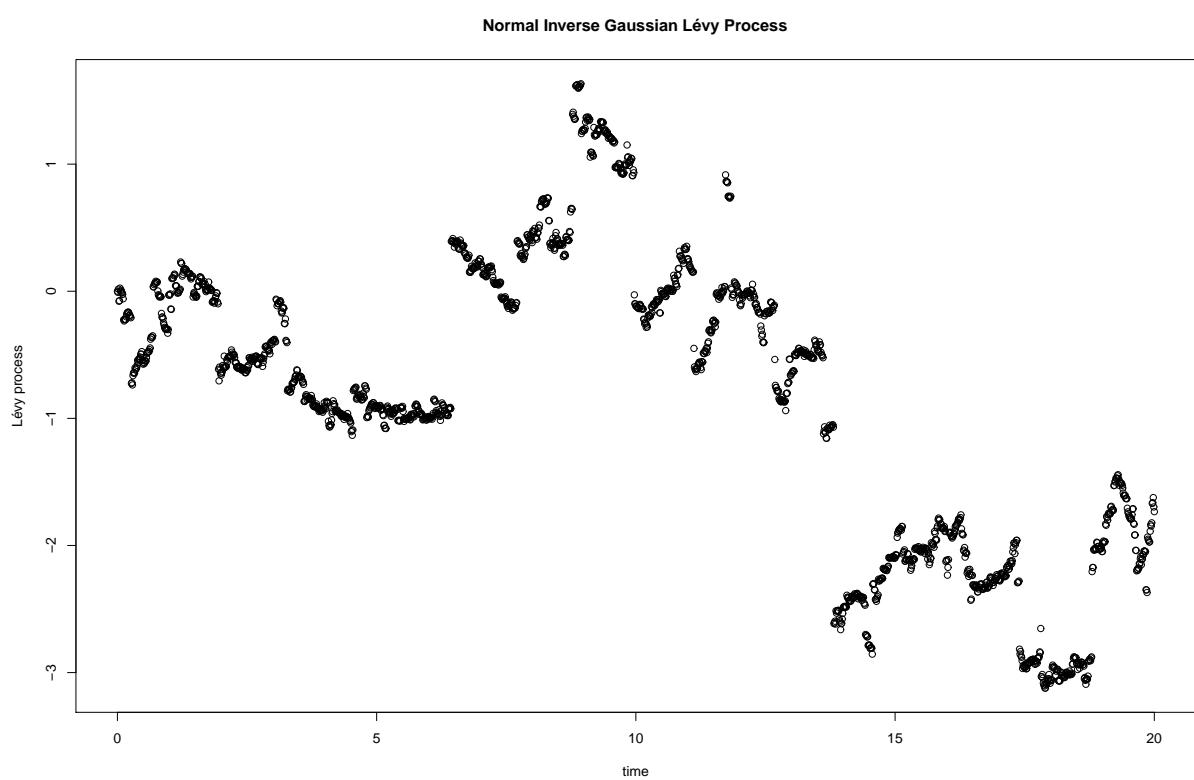


Figure 2: Simulation of a Normal Inverse Gaussian (NIG) Lévy process.

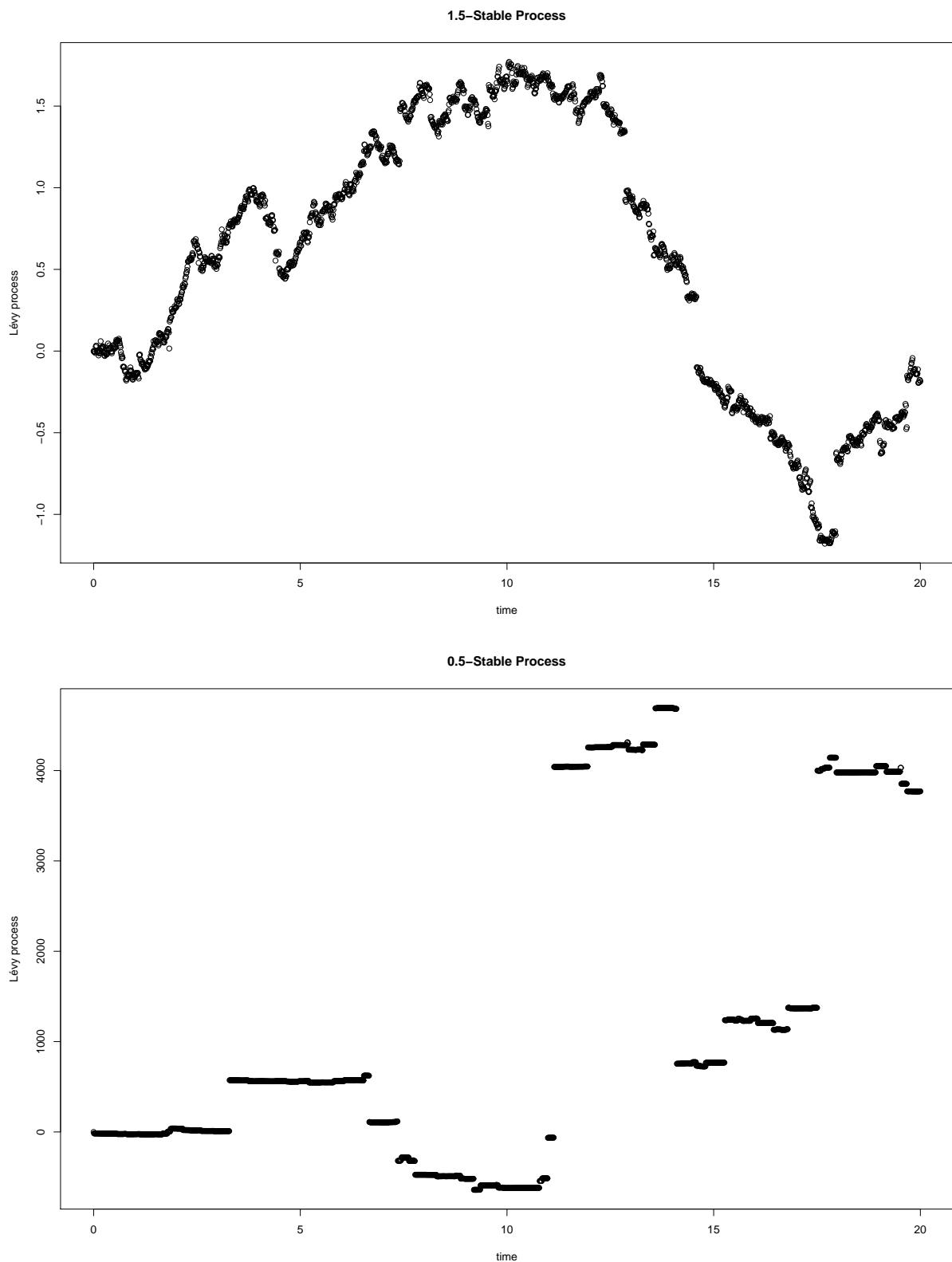


Figure 3: Simulations of stable Lévy processes. A 1.5-stable Lévy process is depicted in the upper row and a 0.5-stable in the lower one.

dexed by  $\mathbb{R}$  not  $\mathbb{R}^+$ . They are obtained by taking two independent copies of a Lévy process and reflecting one copy at the origin.

For detailed expositions on Lévy processes we refer to [1], [6], [37] or [53].

### 3 Definition of CARMA processes and spectral representation

On the intuitive level one wants to be able to interpret a  $d$ -dimensional  $CARMA(p, q)$  process  $Y$  as the stationary solution to the  $p$ -th order linear differential equation

$$P(D)Y_t = (D^p + A_1 D^{p-1} + \dots + A_p)Y_t \quad (3.1)$$

$$= (B_0 D^q + B_1 D^{q-1} + \dots + B_q)DL_t = Q(D)DL_t, \quad (3.2)$$

where the driving input  $L$  is an  $m$ -dimensional Lévy process,  $D$  denotes differentiation with respect to  $t$ , and the coefficients  $A_1, \dots, A_p$  are  $d \times d$  matrices and  $B_0, \dots, B_q$  are  $d \times m$  matrices. The polynomials  $P(z) = z^p + A_1 z^{p-1} + \dots + A_p$  and  $Q(z) = B_0 z^q + B_1 z^{q-1} + \dots + B_q$  with  $z \in \mathbb{C}$  are referred to as the auto-regressive and moving average polynomial, respectively. Finally,  $p, q \in \mathbb{N}$  are the auto-regressive and moving average order.

However, the paths of non-deterministic Lévy processes are not differentiable and so the above equation cannot directly provide a rigorous mathematical definition. Let us briefly consider the case  $(p, q) = (1, 0)$  in which case the resulting process is actually called an Ornstein-Uhlenbeck (OU) process. In the univariate case it is given by the differential equation

$$DY_t = aY_t + DL_t$$

where  $a$  is a real number. So what we basically want is that the change of  $Y$  over an infinitesimal time interval is  $a$  times the current value of the process times the “length of the infinitesimal time interval” plus the change of the Lévy process over the infinitesimal time interval. Rephrasing this idea in the precise language of stochastic differential equations (see e.g. [48]) we obtain

$$dY_t = aY_t dt + dL_t.$$

Using the theory of stochastic differential equations (SDEs) it is easy to see that this SDE has a unique solution given by

$$Y_t = e^{at} Y_0 + e^{at} \int_0^t e^{-as} dL_s.$$

For general orders  $(p, q)$  one could to some extent use a similar reasoning to arrive at a precise definition of CARMA processes. However, we shall take a more elegant route. First note that the differential operators on the auto-regressive side of (3.1) act like integration operators on the moving average side. Hence, they offset the differential operators of the moving average side acting on the Lévy process. Since Lévy processes are not differentiable, we effectively have to integrate at least as often as we differentiate to be able to make sense of (3.1). Hence, a necessary condition ensuring the proper existence of CARMA processes is  $p > q$ .

In order to obtain a rigorous definition of CARMA processes our strategy here shall be to switch from the time domain to the frequency domain where the main tool is the following spectral representation of a Lévy process. Here and in the following we denote by  $A^*$  for a matrix (or vector)  $A$  the Hermitian, i.e. the complex conjugate transposed matrix.

**Theorem 3.1** ([42]). *Let  $(L_t)_{t \in \mathbb{R}}$  be a square integrable  $m$ -dimensional Lévy process with mean  $E[L_1] = 0$  (which implies  $E[L_t] = 0$  for all  $t$ ) and variance  $E[L_1 L_1^*] = \Sigma_L$  at  $t = 1$ . Then there exists a unique  $m$ -dimensional random orthogonal measure  $\Phi_L$  with spectral measure  $F_L$  such that  $E[\Phi_L(\Delta)] = 0$  for any bounded Borel set  $\Delta$ ,  $F_L(dt) = \frac{\Sigma_L}{2\pi} dt$  and*

$$L_t = \int_{-\infty}^{\infty} \frac{e^{i\mu t} - 1}{i\mu} \Phi_L(d\mu), \quad t \in \mathbb{R}.$$

The random measure  $\Phi_L$  is uniquely determined by

$$\Phi_L([a, b)) = \int_{-\infty}^{\infty} \frac{e^{-i\mu a} - e^{-i\mu b}}{2\pi i\mu} dL_\mu \quad (3.3)$$

for all  $-\infty < a < b < \infty$ .

The random orthogonal measure  $\Phi_L$  can intuitively be thought of as the ‘‘Fourier transform’’ of the Lévy process. If  $L_t$  is a Brownian motion, then  $\Phi_L([0, t))$  is again a Brownian motion. For general Lévy processes rather little can be said about the properties of  $\Phi_L$ . For example, it is known that  $\Phi_L$  has second-order stationary and uncorrelated increments, but the increments are neither independent nor stationary in a strict sense, see [27].

In the spectral domain we can now interpret differentiation (and integration) as linear filtering noting that a formal interchange of differentiation and integration gives ‘‘ $DL_t = \int_{-\infty}^{\infty} e^{i\mu t} \Phi_L(d\mu)$ ’’. It can be shown that the resulting process is well-defined whenever the linear filter is square integrable. Thus we obtain as definition for ‘‘ $Y(t) = P(D)^{-1}Q(D)DL(t)$ ’’:

**Definition 3.2** (CARMA Process, [42]). *Let  $L = (L_t)_{t \in \mathbb{R}}$  be a two-sided square integrable  $m$ -dimensional Lévy-process with  $E[L_1] = 0$  and  $E[L_1 L_1^*] = \Sigma_L$ . A  $d$ -dimensional Lévy-driven continuous time autoregressive moving average process  $(Y_t)_{t \in \mathbb{R}}$  of order  $(p, q)$  with  $p, q \in \mathbb{N}_0$  and  $p > q$  (CARMA( $p, q$ ) process) is defined as*

$$\begin{aligned} Y_t &= \int_{-\infty}^{\infty} e^{i\mu t} P(i\mu)^{-1} Q(i\mu) \Phi_L(d\mu), \quad t \in \mathbb{R}, \quad \text{where} \\ P(z) &:= I_m z^p + A_1 z^{p-1} + \dots + A_p, \\ Q(z) &:= B_0 z^q + B_1 z^{q-1} + \dots + B_q \quad \text{and} \end{aligned} \quad (3.4)$$

$\Phi_L$  is the Lévy orthogonal random measure of Theorem 3.1. Here  $A_j \in M_m(\mathbb{R})$ ,  $j = 1, \dots, p$  and  $B_j \in M_{d,m}(\mathbb{R})$  are matrices satisfying  $B_q \neq 0$  and  $\mathcal{N}(P) := \{z \in \mathbb{C} : \det(P(z)) = 0\} \subset \mathbb{R} \setminus \{0\} + i\mathbb{R}$  (i.e. the autoregressive polynomial has no zeros on the complex axis).

Referring to the explicit construction of the random orthogonal measure  $\Phi_L$ , one can easily show that the above defined CARMA processes are necessarily stationary (in the strict sense, i.e. the distributions are left unchanged by a time shift). Since by construction any CARMA process in the sense of Definition 3.2 has a finite variance, it is also weakly stationary, i.e. the second-order moment structure (the variance and autocovariances) are left unchanged by time shifts.

Although the definition of CARMA processes via a spectral representation is elegant and helpful in many theoretical considerations, it is not really usable in applications, as alone simulating a CARMA process from this representation would be a tedious and problematic task. However, luckily we have the following result.

**Theorem 3.3** (State Space Representation, [42]). *Let the Lévy process  $L$  and  $P, Q$  be as before. Define the following coefficient matrices:*

- $\beta_{p-j} = - \sum_{i=1}^{p-j-1} A_i \beta_{p-j-i} + B_{q-j}$ ,  $j = 0, 1, \dots, q$ ,  $\beta_1 = \dots = \beta_{p-q-1} = 0$
- $\beta^* = (\beta_1^*, \beta_2^*, \dots, \beta_p^*)$  and  $A = \left( \begin{array}{c|ccc} 0 & I_{d(p-1)} \\ \hline -A_p & -A_{p-1} & \dots & -A_1 \end{array} \right)$ .

Denote by  $G_t = (G_{1,t}^*, \dots, G_{p,t}^*)^*$  a  $pd$ -dimensional process and assume that  $\mathcal{N}(P) := \{z \in \mathbb{C} : \det(P(z)) = 0\} \subset (-\infty, 0) + i\mathbb{R}$  - the open right half of the complex plane. Then

$$dG_t = AG_t dt + \beta dL_t \quad (3.5)$$

has a unique stationary solution  $G$  given by

$$G_t = \int_{-\infty}^t e^{A(t-s)} \beta dL_s, \quad t \in \mathbb{R}. \quad (3.6)$$

It holds that

$$G_{1,t} = \int_{-\infty}^{\infty} e^{i\mu t} P(i\mu)^{-1} Q(i\mu) \Phi_L(d\mu) = Y_t, \quad t \in \mathbb{R}.$$

So the first  $d$ -components of  $G$  are the CARMA process  $Y$ .

A CARMA process satisfying  $\mathcal{N}(P) := \{z \in \mathbb{C} : \det(P(z)) = 0\} \subset (-\infty, 0) + i\mathbb{R}$  is called causal, because as shown above the value at a time  $t$  only depends on the Lévy process up to time  $t$ , it is a function of  $(L_s)_{s \in (-\infty, t)}$ . In other words a causal CARMA process is fully determined by values in the past. Whenever the condition  $\mathcal{N}(P) := \{z \in \mathbb{C} : \det(P(z)) = 0\} \subset (-\infty, 0) + i\mathbb{R}$  is not satisfied,  $Y_t$  also depends on future values of the Lévy process. In many applications, where it is clear that all we see today can only be influenced by what happened up to now, one only considers causal processes as appropriate models. However there are also applications where non-causal processes are useful. For example, if we want to stochastically model the water level in a river and think of  $t$  as describing the location along the river, both the water levels downstream (in the “future”) and upstream (in the “past”) may influence the water level at a certain point. Note that in this paper we only discuss stationary CARMA processes. In some

applications (e.g. control) it is often adequate to consider non-stationary (non-stable) systems. Then the roots  $\det(P(z)) = 0$  in the set  $(-\infty, 0) + i\mathbb{R}$  describe the stable and causal part of the system and the remaining roots describe the non-stable part.

Theorem 3.3 allows us to treat a causal CARMA process as a solution to the stochastic differential equation (3.5) and thus we can apply all the available results for SDEs. In particular, tasks like simulation of a causal CARMA process are straightforward and easily implemented. However, the above result allows us also to get rid of another restriction. So far we could only define CARMA processes driven by Lévy processes with finite second moments and thus we could so far not have e.g. CARMA processes driven by  $\alpha$ -stable Lévy processes. However, general theory on multidimensional Ornstein-Uhlenbeck processes (see [36] and [54]) tells us that (3.6) is the unique stationary solution to (3.5) as soon as the Lévy process has only a finite logarithmic moment.

**Definition 3.4** (Causal CARMA Process, [42]). *Let  $L = (L_t)_{t \in \mathbb{R}}$  be an  $m$ -dimensional Lévy process satisfying*

$$\int_{\|x\| \geq 1} \ln \|x\| v(dx) < \infty, \quad (3.7)$$

*$p, q \in \mathbb{N}_0$  with  $q < p$ , and further  $A_1, A_2, \dots, A_p \in M_d(\mathbb{R})$ ,  $B_0, B_1, \dots, B_q \in M_{d,m}(\mathbb{R})$ , where  $B_0 \neq 0$ . Define the matrices  $A, \beta$  and the polynomial  $P$  as in Theorem 3.3 and assume  $\sigma(A) = \mathcal{N}(P) \subseteq (-\infty, 0) + i\mathbb{R}$ . Then the  $d$ -dimensional process*

$$Y_t = (I_d, 0, \dots, 0) G_t \quad (3.8)$$

*where  $G_t = \int_{-\infty}^t e^{A(t-s)} \beta dL_s$  is the unique stationary solution to  $dG_t = AG_t dt + \beta dL_t$  is called causal CARMA( $p, q$ ) process.*

*$G$  is referred to as the state space representation.*

A natural question is clearly whether one can also extend the definition of CARMA processes via the spectral representation to the case with infinite variance. For so-called regularly varying Lévy processes with finite mean and thus especially for  $\alpha$ -stable Lévy processes with  $\alpha \in (1, 2)$  a result like Theorem 3.1 has been established in [27]. However, the non-finite variance case is distinctly different, as a limit of integrals has to be taken and the random orthogonal measure is replaced by an object which is – strictly speaking – not even a measure anymore. In that paper a definition of CARMA processes with regularly varying Lévy input analogous to Definition 3.2 has been given and it has been shown that the resulting processes coincide with the causal CARMA processes when both definitions apply. Observe that processes with infinite variance are not only of academic interest, but that they have important applications, for instance, in network data modelling (cf. [43] and [50, 51]). In [29] CARMA processes driven by  $\alpha$ -stable Lévy processes have been successfully used to model electricity prices.

## 4 Properties

In this section we explain and summarise various properties of (causal) CARMA processes.

## 4.1 Second Order Structure

Recall that for convenience we have assumed that the driving Lévy process and thus the CARMA process has mean zero. Looking at the “defining” differential equations, it is clear that if  $E(L_1) = \mu$  then the CARMA process is defined as the one driven by  $L_1 - \mu t$  plus  $A_p^{-1}B_q\mu$  which is then the mean of the CARMA process.

**Proposition 4.1** ([42]). *Let  $Y$  be a (causal) CARMA process driven by a Lévy process  $L$  with finite second moments and set  $\Sigma_L = \text{var}(L_1)$ .*

1. *The CARMA process  $Y$  has autocovariance function:*

$$\text{cov}(Y_{t+h}, Y_t) = \int_{-\infty}^{\infty} \frac{e^{i\mu h}}{2\pi} P(i\mu)^{-1} Q(i\mu) \Sigma_L Q(i\mu)^* (P(i\mu)^{-1})^* d\mu,$$

with  $h \in \mathbb{R}$ .

2. *If  $Y$  is a causal CARMA process, its state space representation  $G$  has the following second order structure:*

$$\begin{aligned} \text{var}(G_t) &= \int_0^{\infty} e^{Au} \beta \Sigma_L \beta^* e^{A^* u} du \\ \text{Avar}(G_t) + \text{var}(G_t) A^* &= -\beta \Sigma_L \beta^* \\ \text{cov}(G_{t+h}, G_t) &= e^{Ah} \text{var}(G_t), h \geq 0. \end{aligned}$$

Since we are only considering stationary CARMA processes, the moments above do not depend on  $t$ .

Since  $Y$  is given by the first  $d$  components of  $G$  the second order structure of  $G$  implies immediately alternative formulae for the second order structure of  $Y$ . In particular, it shows that the autocovariance function always decays like a matrix exponential for  $h \rightarrow \infty$ .

## 4.2 Distribution

Another nice feature is that in principle the distribution of a CARMA process at fixed times as well as the higher dimensional marginal distributions, e.g. the joint distribution of the process at two (or  $n$ ) different points in time, is explicitly known in terms of the characteristic function. The reason is that all these distributions are infinitely divisible and that their Lévy-Khintchine triplet is known in terms of the Lévy-Khintchine triplet of the driving Lévy process. We state this in detail for the stationary distribution in the causal case.

**Proposition 4.2** ([42]). *If  $L$  has characteristic triplet  $(\gamma, \Sigma, \nu)$ , then the stationary distribution of the state space representation  $G$  of a causal CARMA process is infinitely divisible with charac-*

teristic triplet  $(\gamma_G^\infty, \Sigma_G^\infty, v_G^\infty)$ , where

- $\gamma_G^\infty = \int_0^\infty e^{As} \beta \gamma ds + \int_0^\infty \int_{\mathbb{R}^m} e^{As} \beta x [1_{[0,1]}(\|e^{As} \beta x\|) - 1_{[0,1]}(\|x\|)] v(dx) ds,$
- $\Sigma_G^\infty = \int_0^\infty e^{As} \beta \Sigma \beta^* e^{A^* s} ds,$
- $v_G^\infty(B) = \int_0^\infty \int_{\mathbb{R}^m} 1_B(e^{As} \beta x) v(dx) ds$

for all Borel sets  $B \subseteq \mathbb{R}^{pd}$ .

In other words

$$E \left( e^{i\langle u, G_t \rangle} \right) = \exp \left\{ i \langle \gamma_G^\infty, u \rangle - \frac{1}{2} \langle u, \Sigma_G^\infty u \rangle + \int_{\mathbb{R}^{pd}} (e^{i\langle u, x \rangle} - 1 - i \langle u, x \rangle 1_{[0,1]}(\|x\|)) v_G^\infty(dx) \right\}, \quad (4.1)$$

for all  $u \in \mathbb{R}^{pd}$ .

Projection onto the first  $d$  coordinates gives the characteristic triplet of the stationary distribution of  $Y$ . It should, however, be noted that typically the distribution of the CARMA process does not belong to any special family of distributions even if one starts with especially nice Lévy processes.

### 4.3 Dependence Structure

An important property of multivariate stochastic processes, is how their future evolution depends on the past. Suppose that one stands at a given point in time and one disposes of sufficient data at that point to determine the evolution from that point on, also given knowledge of the input from that point onwards. A *Markov process* is a stochastic process, for which the future only depends on the current value and not anymore on the past values (all their information is subsumed in the current value). For a Markov process it – so to speak – only matters where we are now not were we came from. If this characterising property does not only hold at all fixed times, but also at certain random times called stopping times, we speak of a strong Markov process.

**Proposition 4.3** ([42]). *The state space representation  $G$  of a causal CARMA process is a strong Markov process.*

Intuitively it is desirable in many applications that the farther away observations are in time, the less dependent they should be. Usually, one even wants that very far away observations should be basically independent. This idea is mathematically formalized in various concepts of asymptotic independence often referred to as some form of “mixing”.

A comparably weak result which, however, applies to any CARMA process is the following.

**Proposition 4.4** ([28]). *Any stationary CARMA process is mixing.*

Mixing implies ergodicity, i.e. empirically determined moments from the time series converge to the true moments if more and more data is collected. So time averages converge to ensemble averages. This is very important for statistical estimation of CARMA processes, as it implies typically that estimators are consistent (i.e. the estimators converge to the correct value when more and more data is collected).

Typically, one also wants to know the errors of estimators which can be derived from distributional limit results like asymptotic normality. To obtain such results a stronger more uniform notion of asymptotic independence is needed, which is called strong mixing. Typically, one can best establish it for a Markov process.

**Proposition 4.5** ([42]). *For a causal CARMA process with  $E(\|L_1\|^r) < \infty$  for some  $r > 0$  the state space representation  $G$  and the CARMA process  $Y$  are strongly mixing, both with exponentially decaying mixing coefficients.*

## 4.4 Sample Path Properties

Next we look at the sample path properties of a CARMA process.

**Proposition 4.6** ([42]).

- *The sample paths of a CARMA( $p, q$ ) process  $Y$  with  $p > q + 1$  are  $(p - q - 1)$ -times differentiable and for a causal CARMA process it holds that*

$$\frac{d^i}{dt^i} Y_t = G_{i+1,t}, \quad i = 1, 2, \dots, p - q - 1.$$

- *If  $p = q + 1$  and the driving Lévy process has a non-zero Lévy measure  $\nu$  satisfying  $\nu(B_0^{-1}(\mathbb{R}^d \setminus \{0\})) \neq 0$ , then the paths of a CARMA process exhibit jumps and the jumps sizes are given by  $\Delta Y_t := Y_t - Y_{t-} = B_0 \Delta L_t$ .*
- *If the driving Lévy process  $L$  is a Brownian motion, then the sample paths of  $Y$  are continuous and  $(p - q - 1)$ -times continuously differentiable, provided  $p > q + 1$ .*

For examples of the paths of CARMA processes driven by an NIG Lévy process see Figure 4.

## 4.5 Tail Behaviour

As already stated in the introduction one may want to move away from Gaussian models, because extreme (i.e. very low and/or high) observations are far more likely than in a Gaussian distribution. One says that the tails (of the distribution) are heavier than Gaussian ones. Very often it appears also reasonable to use models which are “heavy-tailed” in the sense that only a limited number of moments exists, i.e.  $E(\|X\|^r)$  exists only for low values of  $r$ . Mathematically it is then convenient to use the concept of regular variation (see [22] or [49, 51] for comprehensive introductions in relation to extreme value theory). Roughly speaking this means that the tails behave like a power function when one is far from the centre of the distribution. A random

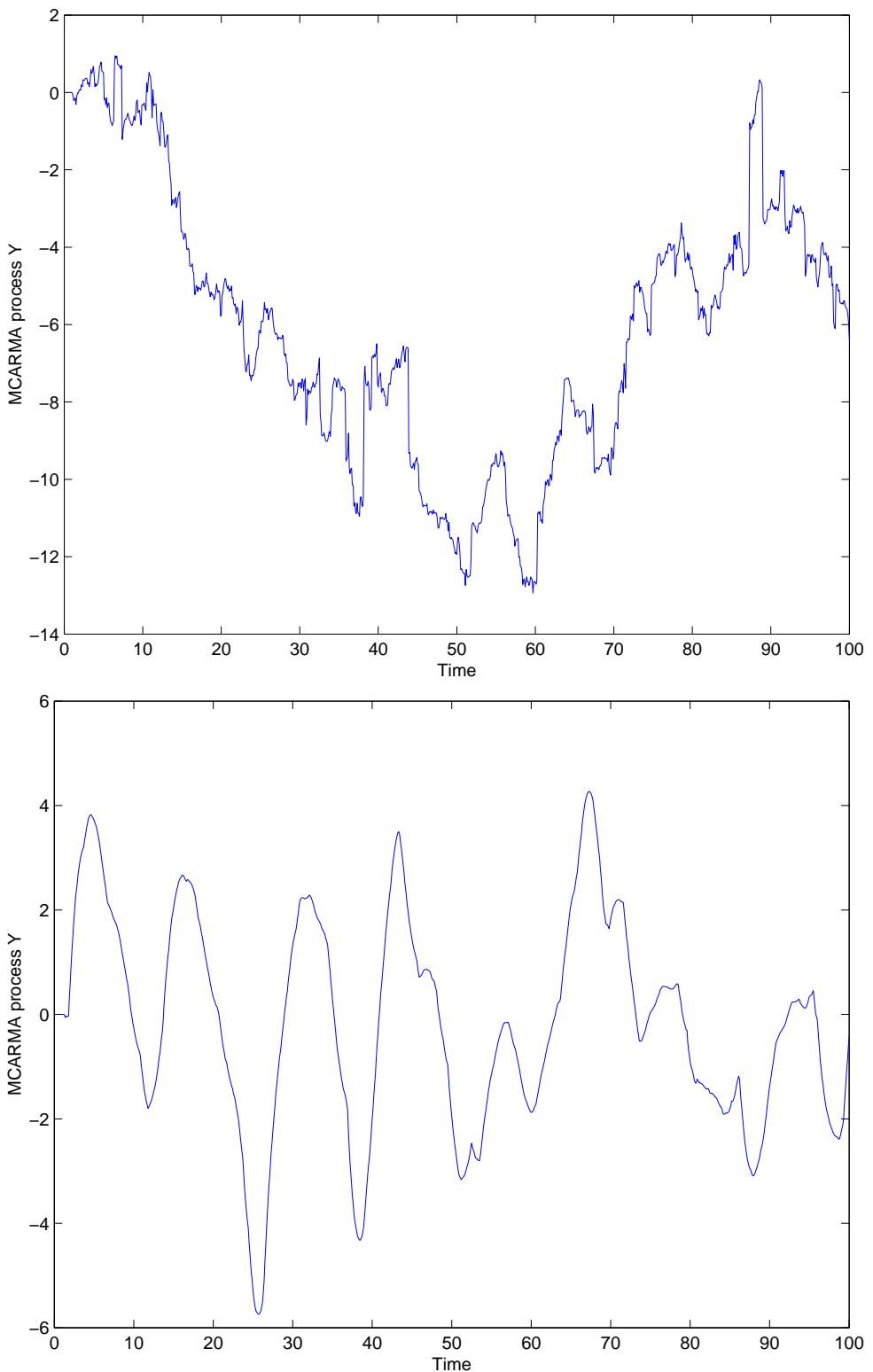


Figure 4: A CARMA(1,0) process driven by an NIG Lévy process having discontinuous paths is shown in the upper display and a CARMA(2,0) process driven by the same Lévy process having continuous paths in the lower one.

variable  $X$  is regularly varying if  $P(\|X\| > x)$  behaves comparably to  $x^{-\alpha}$  for some  $\alpha > 0$  and big values of  $x$ . In [44] (see also [23] in the univariate case) it is shown that under a very mild non-degeneracy condition a CARMA process driven by a regularly varying Lévy process is again regularly varying with the same index  $\alpha$ . Hence, it is straightforward to construct heavy-tailed CARMA processes when applications call for such features.

In the univariate case the tail behaviour of CARMA processes is also understood in certain non-Gaussian situations, where one has lighter tails than regularly varying ones (see [24, 25]).

## 5 State space models

We have defined the causal CARMA process using a so-called state space representation and we have noted that the state space representation  $G$  is made up of the CARMA process  $Y$  and its derivatives as long as they exist. Hence, causal CARMA processes may be viewed as special state space models driven by Lévy processes. In fact, any state space model can also be realized as a CARMA process, as will be shown now.

We start with a precise definition of state space models.

**Definition 5.1.** Let  $\mathbf{L}$  be an  $m$ -dimensional Lévy process and

$$A \in M_N(\mathbb{R}), \quad B \in M_{N,m}(\mathbb{R}), \quad C \in M_{d,N}(\mathbb{R}).$$

A general  $(N,d)$ -dimensional continuous time state space model driven by  $\mathbf{L}$  with parameters  $A, B, C$  is a solution of

$$\begin{aligned} \text{the state equation} \quad d\mathbf{X}_t &= A\mathbf{X}_t dt + Bd\mathbf{L}_t \\ \text{and the observation equation} \quad \mathbf{Y}_t &= C\mathbf{X}_t. \end{aligned}$$

$\mathbf{X}$  is called the state process and  $\mathbf{Y}$  the output process.

Note that the state process is  $N$ -dimensional whereas the output process is  $d$ -dimensional.

Sufficient conditions for the existence of a unique causal stationary solution of the state equation are given by ( $\Re(\cdot)$  indicates the “real part” of a complex number or function)

$$\Re(\lambda_v) < 0, \quad \lambda_v, v = 1, \dots, N, \text{ being the eigenvalues of } A$$

and  $\mathbf{L}$  having finite second moments.

It can easily be shown by integration that  $\mathbf{X}$  satisfies

$$\mathbf{X}_t = e^{A(t-s)}\mathbf{X}_s + \int_s^t e^{A(t-u)}Bd\mathbf{L}_u.$$

Likewise, the stationary output process  $\mathbf{Y}$  satisfies

$$\mathbf{Y}_t = \int_{-\infty}^t Ce^{A(t-u)}Bd\mathbf{L}_u.$$

Its spectral density, the Fourier transform of the autocovariance function, is given by

$$f_{\mathbf{Y}}(\omega) = \frac{1}{2\pi} C(i\omega - A)^{-1} B \Sigma_{\mathbf{L}} B^T (-i\omega - A^T)^{-1} C^T.$$

From Definition 3.4 it is obvious that a CARMA process is a  $(pd, d)$ -dimensional state space model driven by an  $m$ -dimensional Lévy process. The following theorem states that also the converse is true.

**Theorem 5.2 ([55]).** *The stationary solution  $\mathbf{Y}$  of the multivariate state space model  $(A, B, C, \mathbf{L})$  is an  $\mathbf{L}$ -driven CARMA process with autoregressive polynomial  $P$  and moving average polynomial  $Q$  if and only if*

$$C(zI_N - A)^{-1} B = P(z)^{-1} Q(z), \quad \forall z \in \mathbb{C}.$$

*For any  $(A, B, C)$  there exist  $P, Q$  such that the above equation is satisfied and vice versa.*

In reality we typically do not observe some variables of interest continuously, but only at a discrete set of points in time. Let us assume that we sample the process at an equidistant time grid with grid length  $h > 0$  and denote by  $\mathbf{Y}_n^{(h)} := \mathbf{Y}_{nh}$  for  $n \in \mathbb{Z}$  the sampled observations of a state space process.

It is easy to see that

$$\mathbf{Y}_n^{(h)} = C\mathbf{X}_n^{(h)} \tag{5.1}$$

$$\mathbf{X}_n^{(h)} = e^{Ah} \mathbf{X}_n^{(h)} + \int_{(n-1)h}^{nh} e^{A(nh-u)} B d\mathbf{L}_u, \tag{5.2}$$

which immediately shows that  $\mathbf{Y}_n^{(h)}$  is the output process of a discrete time  $(N, d)$ -dimensional state space model driven by the  $N$ -dimensional iid noise  $\left( \int_{(n-1)h}^{nh} e^{A(nh-u)} B d\mathbf{L}_u \right)_{n \in \mathbb{Z}}$ .

It is well-known that any  $(N, d)$ -dimensional state space model in discrete time is an ARMA process. Combining this with Theorem 5.2 tells us that any equidistantly sampled CARMA process  $\mathbf{Y}^{(h)}$  is an ARMA process. This observation will be the basis for estimating CARMA parameters in the next section, where we will need a considerable refinement of this result.

In many applications the sampling frequency is quite high, i.e.  $h$  is very small. Thus it is important to understand how  $\mathbf{Y}^{(h)}$  behaves as  $h \rightarrow 0$  which has been investigated in [18].

As we only observe the process  $\mathbf{Y}$  in a state space model, an important question is what can be said about the state process  $\mathbf{X}$  based on the observations. Hence, we want to reconstruct or “estimate” the latent process  $\mathbf{X}$  as good as possible. This procedure is also referred to as filtering. For Gaussian state space models the easily implementable Kalman filter (see e.g. [11]) is optimal both from a variance point as well as a distributional point of view. For non-Gaussian state space models with finite variance the very same procedure, now typically called linear filtering, gives an “estimate” of the latent process which is the linear (in the observations) “estimate” with the lowest variance. However, it is typically not the “estimate” with the minimal variance and not a conditional expectation. Thus, there are more involved filtering techniques, like particle filtering (see e.g. [20]), which are better.

State space models, mainly Gaussian ones, are also heavily used in stochastic control (see [30] and references therein for a comprehensive overview) and signal processing (see [38, 39, 45], for instance). In both areas one is sometimes dealing with data for which a Gaussianity assumption is not really appropriate due to skewedness, excess kurtosis or heavy-tailedness. Clearly, in such situations Lévy-driven state space models or equivalently CARMA processes should be appealing. Going into the details of the usage in control is beyond the scope of this paper, but it seems worthwhile to mention that there are two uses of state space models in control. Sometimes one assumes that one has some random input which is then “controlled” by the state space model, so the state space model acts as the controller. In contrast to this sometimes the output of the state space model is regarded as the natural output of some system on which an additional controller is acting to ensure that the output meets certain requirements.

## 6 Statistical Estimation

In this section we discuss ways to estimate the parameters of a CARMA process and its driving Lévy process. First we address the estimation of the autoregressive and moving average parameters. Due to parametrisation issues explained later on, we formally do this for Lévy-driven continuous time state space models, as defined in the previous section. In the univariate case quasi-maximum likelihood estimation of CARMA processes is comprehensively studied in [17].

### 6.1 Quasi-maximum likelihood estimation

We assume that we observe the process  $\mathbf{Y}$  at discrete, equally spaced times

$$\mathbf{Y}_n^{(h)} := \mathbf{Y}_{nh}, \quad n \in \mathbb{Z}, \quad h > 0.$$

Furthermore, we define the linear innovations  $\varepsilon^{(h)}$  by

$$\varepsilon_n^{(h)} = \mathbf{Y}_n^{(h)} - P_{n-1} \mathbf{Y}_n^{(h)},$$

where  $P_{n-1}$  denotes the orthogonal projection onto  $\overline{\text{span}} \left\{ \mathbf{Y}_v^{(h)} : -\infty < v < n \right\}$ , i.e. the linear space spanned by the observations until time  $(n-1)h$ . From the construction it is immediate that  $(\varepsilon_n^{(h)})_{n \in \mathbb{Z}}$  is a white noise sequence, i.e. it has mean zero, a constant variance and is uncorrelated. The construction implies that one can only sensibly speak of linear innovations when the driving Lévy process has finite second moments. Thus we will demand the latter for the remainder of this section.

**Theorem 6.1** ([55]). *Assume the eigenvalues  $\lambda_1, \dots, \lambda_N$  of the matrix  $A$  are pairwise distinct and define complex numbers  $\Phi_1, \Phi_2, \dots, \Phi_N$  by*

$$1 - \Phi_1 z - \Phi_2 z^2 - \dots - \Phi_N z^N = \prod_{v=1}^N \left[ 1 - e^{-\lambda_v h} z \right] \quad \forall z \in \mathbb{C}.$$

Then there exist  $\Theta_1, \Theta_2, \dots, \Theta_{N-1}$  in  $M_d(\mathbb{C})$  such that

$$\mathbf{Y}_n^{(h)} - \Phi_1 \mathbf{Y}_{n-1}^{(h)} - \dots - \Phi_N \mathbf{Y}_{n-N}^{(h)} = \varepsilon_n^{(h)} + \Theta_1 \varepsilon_{n-1}^{(h)} + \dots + \Theta_{N-1} \varepsilon_{n-N+1}^{(h)}$$

holds.

Hence,  $\mathbf{Y}^{(h)}$  is a weak ARMA( $N, N-1$ ) process.

This result suggests that one could estimate simply the ARMA coefficients of the sampled process and then transfer these estimates to estimates of the CARMA coefficients. However, to estimate a CARMA process it is not sufficient to estimate an ARMA process, because not all ARMA processes can be embedded in a CARMA process. There are ARMA processes which cannot arise as equidistantly sampled CARMA processes. The way out is carry out the “ARMA estimation” in the CARMA parameter space.

Since we are going to use a quasi-maximum likelihood approach and have discretely sampled observations, all possible models considered in the estimation have to be distinguishable based only on the second-order properties of the sampled process.

**Definition 6.2** (Identifiability). A collection of continuous time stochastic processes  $(\mathbf{Y}_\vartheta, \vartheta \in \Theta)$  is identifiable if for any  $\vartheta_1 \neq \vartheta_2$  the two processes  $\mathbf{Y}_{\vartheta_1}$  and  $\mathbf{Y}_{\vartheta_2}$  have different spectral densities.

It is  $h$ -identifiable,  $h > 0$ , if for any  $\vartheta_1 \neq \vartheta_2$  the two processes  $\mathbf{Y}_{\vartheta_1}^{(h)}$  and  $\mathbf{Y}_{\vartheta_2}^{(h)}$  have different spectral densities.

We assume that our parametrisation is given by a compact parameter space  $\Theta \subset \mathbb{R}^q$  with some  $q \in \mathbb{N}$  and a mapping

$$\psi : \Theta \ni \vartheta \mapsto (A_\vartheta, B_\vartheta, C_\vartheta, \mathbf{L}_\vartheta).$$

Here,  $A_\vartheta$  is the  $N \times N$  matrix of our Definition 5.1 dependent on the parameters  $\vartheta$  and likewise for  $B_\vartheta, C_\vartheta$  and  $\mathbf{L}_\vartheta$ .

We need to ensure that our parametrisation is minimal regarding the dimensions, since a fixed output process can result from artificially arbitrarily high-dimensional state space models.

**Assumption P1** (Minimality). For all  $\vartheta \in \Theta$  the triple  $(A_\vartheta, B_\vartheta, C_\vartheta)$  is minimal in the sense that if

$$C(zI_m - A)^{-1}B = C_\vartheta(zI_N - A_\vartheta)^{-1}B_\vartheta$$

then  $m \geq N$  must be true.

**Assumption P2** (Eigenvalues). For all  $\vartheta \in \Theta$  the eigenvalues of  $A_\vartheta$  are pairwise distinct and contained in the strip

$$\{z \in \mathbb{C} : -\pi/h < \Im(z) < \pi/h\}.$$

We want to use a parametrisation for the continuous time state space model, but need to ensure that it is  $h$ -identifiable. The following theorem provides easy-to-check criteria.

**Theorem 6.3** ([56]). Assume that the parametrisation  $\psi : \Theta \ni \vartheta \mapsto (A_\vartheta, B_\vartheta, C_\vartheta, \mathbf{L}_\vartheta)$  is

- *identifiable*,
- *minimal*
- *and satisfies the eigenvalue condition.*

Then the corresponding collection of output processes  $\{\mathbf{Y}_\vartheta, \vartheta \in \Theta\}$  is  $h$ -identifiable.

The quasi-maximum likelihood (QML) estimator is now obtained by pretending the observations were Gaussian, taking the corresponding likelihood and maximising it. More precisely the QML of  $\vartheta$  based on  $L$  observations  $\mathbf{y}^L = (\mathbf{y}_1, \dots, \mathbf{y}_L)$  (of a CARMA process with parameter  $\vartheta_0$ ) is

$$\hat{\vartheta}^L = \operatorname{argmax}_{\vartheta \in \Theta} \mathcal{L}_\vartheta(\mathbf{y}^L),$$

where  $\mathcal{L}_\vartheta$  is the Gaussian likelihood function which is proportional to

$$\left( \prod_{n=1}^L \det V_{\vartheta,n} \right)^{-1/2} \exp \left\{ -\frac{1}{2} \sum_{n=1}^L \mathbf{e}_{\vartheta,n}^T V_{\vartheta,n}^{-1} \mathbf{e}_{\vartheta,n} \right\}$$

with

$$\begin{aligned} \mathbf{e}_{\vartheta,n} &= \mathbf{y}_n - P_{n-1} \mathbf{Y}_{\vartheta,n}^{(h)} \Big|_{\{\mathbf{Y}_{\vartheta,v}^{(h)} = \mathbf{y}_v : 1 \leq v < n\}}, \\ V_{\vartheta,n} &= \mathbb{E} \left[ \mathbf{e}_{\vartheta,n} \mathbf{e}_{\vartheta,n}^T \Big| \mathbf{Y}_{\vartheta,v}^{(h)} = \mathbf{y}_v : 1 \leq v < n \right]. \end{aligned}$$

So  $\mathbf{e}_{\vartheta,n}$  are the linear innovations under the model given by  $\vartheta$  and  $V_{\vartheta,n}$  are their variances or the one-step prediction errors. Note that  $\mathbf{y}^L$  are in contrast to this observations of the CARMA process with the unknown parameter  $\vartheta_0$  which we are about to estimate.

Computing the QML estimator is now a straightforward task utilising the Kalman recursions and numerically maximising the likelihood. However, since we have not used the true likelihood, it is not clear whether the resulting estimators are really sensible in the sense that they converge to the true parameters. Luckily, one can show that the estimators are well-behaved.

**Theorem 6.4** (Strong consistency, [56]). *Assume the parametrisation  $\psi$  is continuous. For every sampling interval  $h > 0$ , the QML estimator  $\hat{\vartheta}^L$  is strongly consistent, i.e.*

$$\hat{\vartheta}^L \rightarrow \vartheta_0 \quad \text{a.s. as } L \rightarrow \infty,$$

*provided the parametrisation is  $h$ -identifiable.*

However, so far we cannot assess the quality of our estimators by confidence intervals etc., which is made possible by the following result.

**Theorem 6.5** (Asymptotic normality, [56]). *Assume that the driving Lévy process satisfies  $E||\mathbf{L}_{\vartheta_0}(1)||^{4+\delta} < \infty$  for some  $\delta > 0$  and that the parametrisation  $\psi$  is three times continuously differentiable. For every sampling interval  $h > 0$ , the QML estimator  $\hat{\vartheta}^L$  is asymptotically normally distributed, i.e.*

$$\sqrt{L}(\hat{\vartheta}^L - \vartheta_0) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Omega), \quad \Omega = J(\vartheta_0)^{-1} I(\vartheta_0) J(\vartheta_0)^{-1},$$

with

$$J(\vartheta) = \lim_{L \rightarrow \infty} \frac{1}{L} \frac{\partial^2}{\partial \vartheta \partial \vartheta^T} \ln \mathcal{L}_\vartheta(\mathbf{y}^L),$$

$$I(\vartheta) = \lim_{L \rightarrow \infty} \frac{1}{L} \text{Var} \frac{\partial}{\partial \vartheta} \ln \mathcal{L}_\vartheta(\mathbf{y}^L),$$

provided the parametrisation is  $h$ -identifiable.

To obtain identifiable parametrisations one uses like in the discrete time case (see [32] or [40], for instance) so called canonical parametrisations like the echelon state space form. For more details on this we refer to [56]. Since such parametrisations are typically available for state space models rather than CARMA processes, one normally estimates state space models rather than the equivalent CARMA processes.

Let us finally look at one simulation study.

A  $d$ -dimensional normal inverse Gaussian (NIG) Lévy process  $\mathbf{L}$  (see e.g. [3, 7, 46]) with parameters

$$\delta > 0, \kappa > 0, \beta \in \mathbb{R}^d, \Delta \in M_d^+(\mathbb{R})$$

is given by a normal mean-variance mixture, i.e.

$$\mathbf{L}_t = \mu + V\Delta\beta + V^{1/2}N,$$

where  $N$  is  $d$ -dimensionally normally distributed with mean zero and variance  $\Delta$  and independent of

$$V \sim \text{IG}(\delta/\kappa, \delta^2)$$

which follows a so-called inverse Gaussian distribution ([35]).

We consider now a bivariate NIG-driven CARMA process with zero mean given by the state space form

$$d\mathbf{X}_t = \begin{bmatrix} \vartheta_1 & \vartheta_2 & 0 \\ 0 & 0 & 1 \\ \vartheta_3 & \vartheta_4 & \vartheta_5 \end{bmatrix} \mathbf{X}_t dt + \begin{bmatrix} \vartheta_1 & \vartheta_2 \\ \vartheta_6 & \vartheta_7 \\ \vartheta_3 + \vartheta_5 \vartheta_6 & \vartheta_4 + \vartheta_5 \vartheta_7 \end{bmatrix} d\mathbf{L}_t,$$

$$\mathbf{Y}_t = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \mathbf{X}_t, \quad \Sigma_{\mathbf{L}} = \begin{bmatrix} \vartheta_8 & \vartheta_9 \\ \vartheta_9 & \vartheta_{10} \end{bmatrix}.$$

The parameters are  $\vartheta_1, \vartheta_2, \dots, \vartheta_{10}$  and the parametrisation is in one of the canonical identifiable forms.

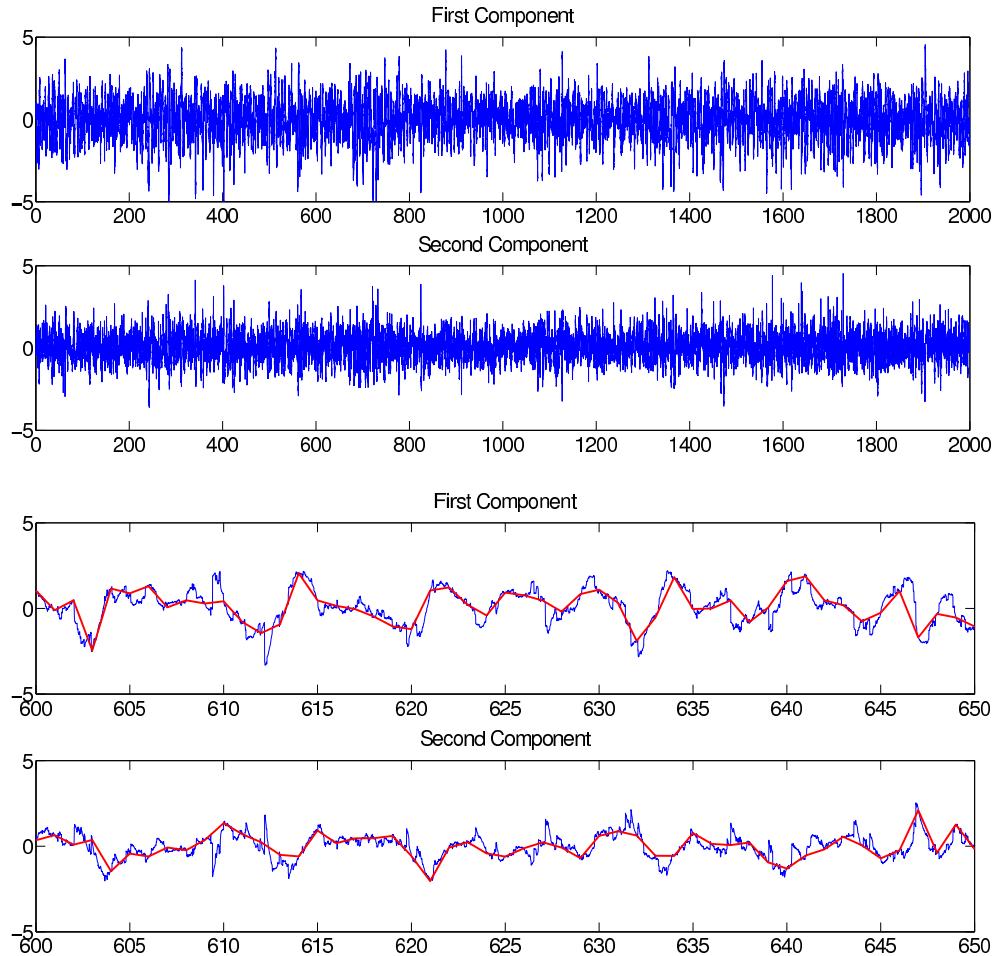


Figure 5: One realisation of a bivariate NIG-driven CARMA process (upper two displays) and the effect of sampling (lower two displays). The linearly interpolated process over the time interval  $[600, 650]$  resulting from sampling at integer times is shown as the thicker line, whereas the thinner line is the true CARMA process.

parameter	sample mean	sample bias	sample standard deviation	estimated standard deviation
$\vartheta_1$	-1.0001	0.0001	0.0354	0.0381
$\vartheta_2$	-2.0078	0.0078	0.0479	0.0539
$\vartheta_3$	1.0051	-0.0051	0.1276	0.1321
$\vartheta_4$	-2.0068	0.0068	0.1009	0.1202
$\vartheta_5$	-2.9988	-0.0012	0.1587	0.1820
$\vartheta_6$	1.0255	-0.0255	0.1285	0.1382
$\vartheta_7$	2.0023	-0.0023	0.0987	0.1061
$\vartheta_8$	0.4723	-0.0028	0.0457	0.0517
$\vartheta_9$	-0.1654	0.0032	0.0306	0.0346
$\vartheta_{10}$	0.3732	0.0024	0.0286	0.0378

Table 1: Summary of the results of the simulation study on the QML estimation of a bivariate NIG-driven CARMA process. The second column states the mean of estimators obtained over 350 simulated paths, the third column the resulting bias and the fourth column the standard deviation of the obtained estimators. Finally, the last column states the standard deviation for the estimators as predicted by the asymptotic normality result Theorem 6.5.

A simulated path is shown in Figure 5.

We calculated the QML estimates for this bivariate NIG-driven CARMA process based on observations over the time horizon  $[0, 2000]$  at integer times and repeated this for 350 different simulated paths. The estimation results are summarised in Table 1. It shows that the sample bias of the obtained estimators in the simulation study is very small and that the sample standard deviation is close to the standard deviation predicted by the asymptotic normality result Theorem 6.5. Actually, the sample standard deviation is always smaller which is nice, as it implies that the standard deviation predicted by the asymptotic normality result Theorem 6.5 is a conservative estimate.

## 6.2 Statistical inference for the driving Lévy process

The above quasi-maximum likelihood approach only allows to estimate the autoregressive and moving average parameters as well as the variance of the driving Lévy process. However, typically we want to estimate many more parameters of the driving Lévy process or even first need to get an idea to which family the driving Lévy process may belong to. To this end one can reconstruct from the CARMA process the driving Lévy process. Typically, the CARMA process is only observed at a discrete set of times and then the best we can do is to get approximations of the increments of the Lévy process. One can then treat the approximate increments as if they were the true ones of the Lévy process. ‘‘Looking’’ at them one should be able to choose appropriate parametric families. By using the approximate increments, as one would use the true ones, in maximum likelihood or method of moment based estimation procedures one can do parametric

inference for the Lévy process. The construction of the approximate increments and their use in estimation procedures has been studied in detail in [14] where it is in particular shown that the estimators are good in the sense that they are consistent and asymptotically normal under reasonable assumptions when taking appropriate limits.

It should be noted that the idea to reconstruct the Lévy process can already be found in [10] or [17]. In the following we illustrate this approach for a univariate Ornstein-Uhlenbeck, i.e. a CARMA(1,0), process based on [16] and [31] from which all examples and plots are taken.

Recall that an Ornstein-Uhlenbeck (OU) process is the unique strictly stationary solution to

$$dY_t = aY_t dt + dL_t. \quad (6.1)$$

where  $(L_t)_{t \in \mathbb{R}}$  is a Lévy process with  $E(\ln(\max(|L_1|, 1))) < \infty$  and autoregressive parameter  $a < 0$ . The solution of the stochastic differential equation is given explicitly by

$$Y_t = e^{a(t-s)} Y_s + \int_s^t e^{a(t-u)} dL_u. \quad (6.2)$$

If the OU process is observed continuously on  $[0, T]$ , then the integrated form of (6.1) immediately gives

$$L_t = Y_t - Y_0 - a \int_0^t Y_s ds.$$

The increments of the driving Lévy process  $\Delta L_n^{(h)}$  on the intervals  $((n-1)h, nh]$  with  $n \in \mathbb{N}$  can be represented as

$$\Delta L_n^{(h)} := L_{nh} - L_{(n-1)h} = Y_{nh} - Y_{(n-1)h} - a \int_{(n-1)h}^{nh} Y_u du. \quad (6.3)$$

What we want, is to approximately reconstruct the sequence  $\Delta L^{(h)}$  of increments over intervals of length  $h$  from observations of the CARMA process made over a finer equidistant grid. To this end one simply approximates the integral  $\int_{(n-1)h}^{nh} Y_u du$  by some numerical integration scheme needing only the values of the process on this finer grid. Since the approximations of  $\Delta L^{(h)}$  become thus closer and closer to the true increment as the numerical integration scheme becomes more exact, [14] derive their asymptotic results when both the observation interval as well as the observation frequency goes to infinity. Note that in practice one does not know  $a$  so one has to estimate it first, which could e.g. be done by the already described quasi-maximum likelihood approach.

Turning to an example, let us consider the OU process given by

$$dX_t = -0.6X_t dt + dL_t, \quad (6.4)$$

with  $L$  being a standardised Gamma process, i.e.  $L_t$  has density

$$f_{L_t}(x) = \frac{\gamma^{1/2\gamma t}}{\Gamma(\gamma t)} x^{\gamma t - 1} e^{-x\gamma^{1/2}} \mathbf{1}_{[0, \infty)},$$

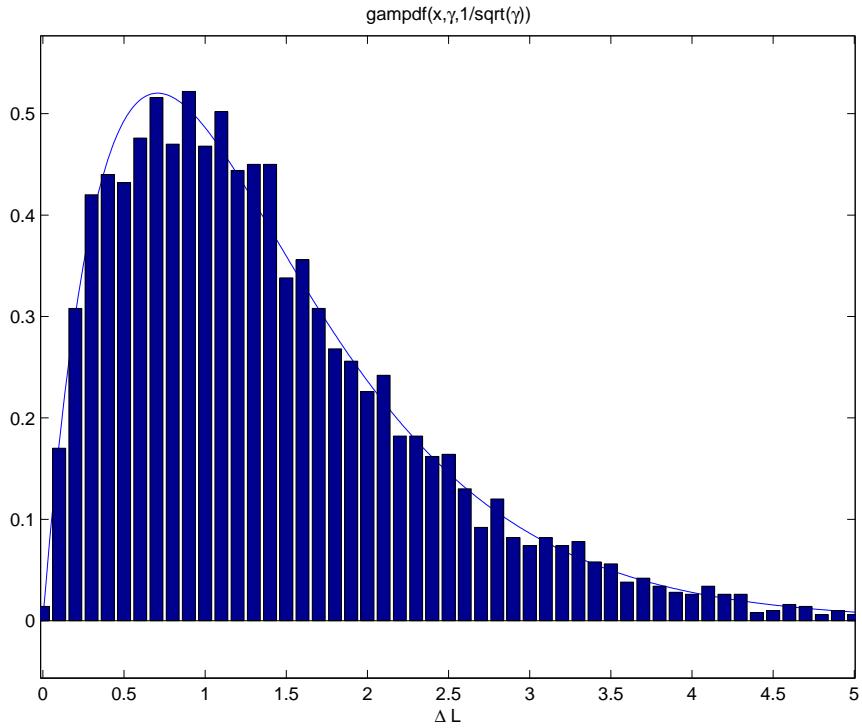


Figure 6: Probability density of the increments of the standardised Lévy process with  $\gamma = 2$  and the histogram of the estimated increments from one path of the OU process, obtained by sampling the process with grid length 0.01. (Source: [31].)

and the parameter  $\gamma$  being set to 2.

In [31] 100 paths of this OU process on the time interval  $[0, 5000]$  have been simulated and then the Lévy increments over time intervals of unit length have been approximated by sampling the OU process over a grid of size  $h$ .

In Figure 6 the histogram of the Lévy increments distribution from one path with  $h = 0.01$  is shown, together with the true probability density of  $L_1$ .

If one further averages over all one hundred paths which is equivalent to looking at one path over a one hundred times longer time horizon, the fit of the histogram to the true density becomes visually almost perfect, see Figure 7.

Based on the approximate Lévy increments one can now estimate the parameter  $\gamma$  by maximum likelihood. Table 2 shows summary statistics of the resulting estimator for different sampling grid sizes  $h$ . The data in the table is based on estimating  $\gamma$  separately for each of the 100 simulated paths.

To conclude, the simulation study illustrates that the recovery of the background driving Lévy process and the parametric estimation based on the approximate increments works quite well.

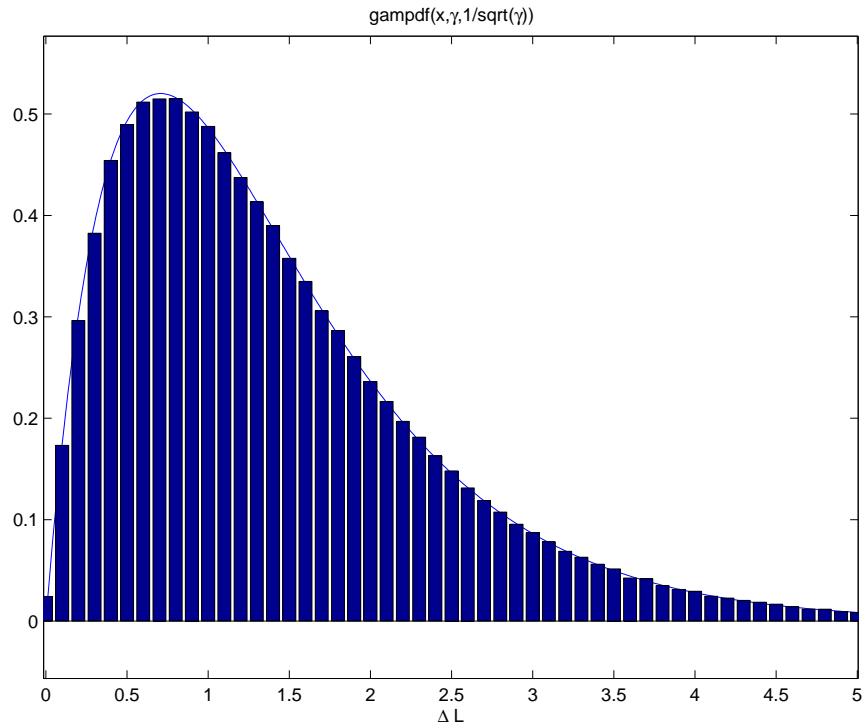


Figure 7: Probability density of the increments of the standardised Lévy process with  $\gamma = 2$  and the histogram of the estimated increments for all 100 paths of the OU process, obtained by sampling the process with grid length 0.01. (Source: [31].)

Table 2: Estimated parameters of the standardised driving Lévy process based on 100 paths on  $[0, 5000]$  of the Gamma-driven OU process.

$h$	Parameter	Sample mean of estimator	Sample standard deviation of estimator
0.01	$\gamma$	2.0039	0.0314
0.1	$\gamma$	2.0043	0.0340
1	$\gamma$	1.9967	0.0539

## 7 Concluding Remarks

Finally, we would like to mention that there are other stochastic models like the so-called ECO-GARCH process of [33] and [34] where CARMA processes are an important ingredient as well as extensions of CARMA processes. One extension are fractionally integrated CARMA (FICARMA) processes (see [13] and [41]). While CARMA processes have an exponentially decaying autocovariance function and thus have always short memory, FICARMA processes exhibit polynomially decaying autocovariance functions and are thus able to model long memory phenomena (see [21] or [52] for detailed introductions into the topic of long range dependence). However, the paths of FICARMA processes are continuous. A class of processes with possible long memory, jumps in the paths and related to CARMA processes are the supOU processes, see [4, 5] and [26]. As noted in [5] multivariate supOU processes can be straightforwardly extended to obtain so-called supCARMA processes. Long memory is (believed to be) encountered in data from many different areas, e.g. finance or telecommunication. Since it is an asymptotic property and similar effects in the autocorrelation function might be caused by structural breaks (non-stationarity), it is often hardly debated whether there truly is long memory in a time series. The first scientific study considering long range dependence properties was looking at the water level of the river Nile (see [57]).

From the overview on CARMA processes presented in this paper it should not only be clear that they are useful in many applications, but also that there are still many questions to be addressed in future research. These include alternative estimators to the ones presented here, estimators which work in the heavy-tailed case when one does not have a finite variance or order selection, i.e. a theory how to choose the orders  $(p, q)$  of the autoregressive and moving average polynomial when one fits CARMA processes to observed time series.

## Acknowledgements

The author takes pleasure in thanking Florian Fuchs, Claudia Klüppelberg and Eckhard Schlemm for comments on previous drafts and Maria Graf for allowing him to use material from her diploma thesis in Section 6.2. Financial support from the TUM Institute for Advanced Study funded by the German Excellence Initiative through a Carl-von-Linde Junior Fellowship is gratefully acknowledged.

## References

- [1] D. Applebaum. *Lévy Processes and Stochastic Calculus*, volume 93 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2004.
- [2] M. Arató. *Linear Stochastic Systems with Constant Coefficients*, volume 45 of *Lecture Notes in Control and Information Sciences*. Springer, Berlin, 1982.

- [3] O. E. Barndorff-Nielsen. Processes of normal inverse Gaussian type. *Finance Stoch.*, 2: 41–68, 1998.
- [4] O. E. Barndorff-Nielsen. Superposition of Ornstein–Uhlenbeck type processes. *Theory Probab. Appl.*, 45:175–194, 2001.
- [5] O. E. Barndorff-Nielsen and R. Stelzer. Multivariate supOU processes. *Ann. Appl. Probab.*, 21:140–182, 2011.
- [6] J. Bertoin. *Lévy Processes*. Cambridge University Press, Cambridge, 1998.
- [7] P. Blæsild and J. L. Jensen. Multivariate distributions of hyperbolic type. In C. Taillie, G. P. Patil, and B. A. Baldessari, editors, *Statistical distributions in scientific work - Proceedings of the NATO Advanced Study Institute held at the Université degli Studi di Trieste, Triest, Italy, July 10 - August 1, 1980*, volume 4, pages 45–66, Dordrecht, Holland, 1981. D. Reidel.
- [8] P. J. Brockwell. Lévy-driven CARMA processes. *Ann. Inst. Statist. Math.*, 52:1–18, 2001.
- [9] P. J. Brockwell. Continuous-time ARMA processes. In D. N. Shanbhag and C. R. Rao, editors, *Stochastic Processes: Theory and Methods*, volume 19 of *Handbook of Statistics*, pages 249–276, Amsterdam, 2001. Elsevier.
- [10] P. J. Brockwell. Lévy-driven continuous-time ARMA processes. In T. G. Andersen, R. Davis, J.-P. Kreiß, and T. Mikosch, editors, *Handbook of Financial Time Series*, pages 457 – 480, Berlin, 2009. Springer.
- [11] P. J. Brockwell and R. A. Davis. *Time Series: Theory and Methods*. Springer, New York, 2nd edition, 1991.
- [12] P. J. Brockwell and A. Lindner. Existence and uniqueness of stationary Lévy-driven CARMA processes. *Stoch. Process. Appl.*, 119:2660–2681, 2009.
- [13] P. J. Brockwell and T. Marquardt. Lévy-driven and fractionally integrated ARMA processes with continuous time parameter. *Statist. Sinica*, 15:477–494, 2005.
- [14] P. J. Brockwell and E. Schlemm. Parametric estimation of the driving Lévy process of multivariate CARMA processes from discrete observations. submitted for publication, 2011.
- [15] P. J. Brockwell, R. A. Davis, and Y. Yang. Continuous-time Gaussian autoregression. *Statist. Sinica*, 17:63–80, 2007.
- [16] P. J. Brockwell, R. A. Davis, and Y. Yang. Estimation for nonnegative Lévy-driven Ornstein-Uhlenbeck processes. *J. Appl. Probab.*, 44:977–989, 2007.
- [17] P. J. Brockwell, R. A. Davis, and Y. Yang. Estimation of non-negative Lévy-driven CARMA processes. *J. Bus. Econom. Statist.*, 2011. to appear.

- [18] P. J. Brockwell, V. Ferrazzano, and C. Klüppelberg. High frequency sampling of a continuous-time ARMA process. submitted for publication, 2011. available at: <http://www-m4.ma.tum.de>.
- [19] J. L. Doob. The elementary Gaussian processes. *Ann. Math. Stat.*, 25:229–282, 1944.
- [20] A. Doucet, N. De Freitas, and N. J. Gordon, editors. *Sequential Monte Carlo Methods in Practice*, Statistics for Engineering and Information Science, New York, 2001. Springer.
- [21] P. Doukhan, M. S. Taqqu, and G. Oppenheim, editors. *Long-Range Dependence*, Boston, 2003. Birkhäuser.
- [22] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin, 1997.
- [23] V. Fasen. Extremes of regularly varying Lévy-driven mixed moving average processes. *Adv. in Appl. Probab.*, 37:993–1014, 2005.
- [24] V. Fasen. Extremes of subexponential Lévy driven moving average processes. *Stoch. Proc. Appl.*, 116:1066–1087, 2006.
- [25] V. Fasen. Extremes of Lévy driven mixed MA processes with convolution equivalent distributions. *Extremes*, 12:265–296, 2009.
- [26] V. Fasen and C. Klüppelberg. Extremes of supOU processes. In F. E. Benth, G. Di Nunno, T. Lindstrom, B. Øksendal, and T. Zhang, editors, *Stochastic Analysis and Applications: The Abel Symposium 2005*, volume 2 of *Abel Symposia*, pages 340–359, Berlin, 2007. Springer.
- [27] F. Fuchs and R. Stelzer. Spectral representation of multivariate regularly varying Lévy and CARMA processes. *J. Theoret. Probab.*, 2011. to appear.
- [28] F. Fuchs and R. Stelzer. Mixing conditions for multivariate infinitely divisible processes with an application to mixed moving averages and the supOU stochastic volatility model. *ESAIM: Probab. Stat.*, 2011. to appear.
- [29] I. García, C. Klüppelberg, and G. Müller. Estimation of stable CARMA models with an application to electricity spot prices. *Stat. Model.*, 2010. to appear.
- [30] H. Garnier and L. Wang, editors. *Identification of Continuous-time Models from Sampled Data*. Advances in Industrial Control. Springer, London, 2008.
- [31] M. Graf. Parametric and nonparametric estimation of positive Ornstein-Uhlenbeck type processes. Diploma thesis, Centre for Mathematical Sciences, TU München, 2009. Available from <http://www-m4.ma.tum.de/Diplarb>.

- [32] E. J. Hannan and M. Deistler. *The statistical theory of linear systems*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, New York, 1988. ISBN 0-471-80777-X.
- [33] S. Haug and C. Czado. An exponential continuous time GARCH process. *J. Appl. Probab.*, 44:960–976, 2007.
- [34] S. Haug and R. Stelzer. Multivariate ECOGARCH processes. *Econometric Theory*, 2010. accepted for publication.
- [35] B. Jørgensen. *Statistical Properties of the Generalized Inverse Gaussian Distribution*. Lecture Notes in Statistics. Springer, New York, 1982.
- [36] Z. J. Jurek and D. J. Mason. *Operator-limit Distributions in Probability Theory*. John Wiley & Sons, New York, 1993.
- [37] A. E. Kyprianou. *Introductory Lectures on Fluctuations of Lévy Processes with Applications*. Universitext. Springer, Berlin, 2006.
- [38] E. K. Larsson. Limiting sampling results for continuous-time ARMA systems. *Internat. J. Control*, 78:461–473, 2005.
- [39] E. K. Larsson, M. Mossberg, and T. Söderström. An overview of important practical aspects of continuous-time ARMA system identification. *Circuits Systems Signal Process.*, 25:17–46, 2006.
- [40] H. Lütkepohl. *New Introduction to Multiple Time Series Analysis*. Springer, Berlin, 2005.
- [41] T. Marquardt. Multivariate fractionally integrated CARMA processes. *J. Multivariate Anal.*, 98:1705–1725, 2007.
- [42] T. Marquardt and R. Stelzer. Multivariate CARMA processes. *Stochastic Process. Appl.*, 117:96–120, 2007.
- [43] T. Mikosch, S. I. Resnick, R. Holger, and A. Stegeman. Is network traffic approximated by stable Lévy motion or fractional Brownian motion? *Ann. Appl. Probab.*, 12:23–68, 2002.
- [44] M. Moser and R. Stelzer. Tail Behavior of Multivariate Lévy-Driven Mixed Moving Average Processes and supOU Stochastic Volatility Models. *Adv. in Appl. Probab.*, 43:1109–1135, 2011.
- [45] M. Mossberg and E. K. Larsson. Fast and approximative estimation of continuous-time stochastic signals from discrete-time data. In *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04)*, volume 2, pages 529–532, Montreal, QC, May 17-21, 2004, 2004.

- [46] K. Prause. *The Generalized Hyperbolic Model: Estimation, Financial Derivatives and Risk Measures*. Dissertation, Mathematische Fakultät, Albert-Ludwigs-Universität Freiburg i. Br., Freiburg, Germany, 1999.
- [47] M. B. Priestley. *Spectral analysis and time series. Vol. 1 & 2*. Academic Press, London, 1981.
- [48] P. Protter. *Stochastic Integration and Differential Equations*, volume 21 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, New York, 2nd edition, 2004.
- [49] S. I. Resnick. *Extreme Values, Regular Variation and Point Processes*. Springer, New York, 1987.
- [50] S. I. Resnick. Heavy tail modeling and teletraffic data: special invited paper. *Ann. Statist.*, 25:1805–1869, 1997.
- [51] S. I. Resnick. *Heavy-Tail Phenomena*. Springer, 2007.
- [52] G. Samorodnitsky. Long range dependence. *Found. Trends Stoch. Syst.*, 1:163–257, 2006.
- [53] K. Sato. *Lévy Processes and Infinitely Divisible Distributions*, volume 68 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999.
- [54] K. Sato and M. Yamazato. Operator-selfdecomposable distributions as limit distributions of processes of Ornstein-Uhlenbeck type. *Stochastic Process. Appl.*, 17:73–100, 1984.
- [55] E. Schlemm and R. Stelzer. Multivariate CARMA processes, continuous-time state space models and complete regularity of the innovations of the sampled processes. *Bernoulli*, 2011. accepted for publication.
- [56] E. Schlemm and R. Stelzer. Quasi maximum likelihood estimation for multivariate Lévy driven CARMA processes. submitted for publication, 2011.
- [57] O. Tousson. Mémoire sur l’Histoire du Nil. *Mémoires de l’Institut d’Égypte (Cairo)*, 1925.